

• 外国哲学 •

为他心辩护^{*}

——处理他心问题的一种复合方案

王 晓 阳

〔摘 要〕“他心问题”是近代以来知识论的一个重要话题。它关注的是他人或他物是否存在与我相类似的心灵状态或意识状态的问题。处理他心问题的两大当代基础理论方案分别是推论主义和非推论主义。通过分析这两大基础理论方案可以发现，在应对关于他心的怀疑论时，推论主义无法应对极端版本的怀疑论，非推论主义则无法应对温和版本的怀疑论。而一旦我们按照某种方式把推论主义与非推论主义各自的优势相结合，那么就可以构造出一个处理他心问题的新方案——复合方案。本文所提供的这样一种复合方案不但可以有效应对来自（温和的与极端的）两个版本怀疑论的挑战，而且可以为我们提供一个关于他心的完整理解。

〔关键词〕他心问题 推论主义 非推论主义 复合方案 知识论

〔中图分类号〕B565.6

一、引言：他心问题与怀疑论

在由笛卡尔开启的西方知识论传统中，他心问题（the Problem of Other Minds）是个经久不衰的重要议题。它也是人类一直以来试图攻克科学难关之一。一般来说，他心问题关注他者的心灵状态（mental states）或意识状况（conscious status）^①，因而可以将他心问题细分为如下两个相关问题：

问题 1：除我之外，他人或他物有心灵或意识么？

问题 2：如果有，是否与我的心灵状态或意识状况类似？

不难理解，对问题 1 和问题 2 的回答，要视情况而定。具体而言，如果考虑他人，那么，对于我们中的大多数而言，问题 1 和问题 2 的答案显然都是肯定的。如果考虑他物（即非人对象），情况会复杂一些：如果某些广为人知的科学理论（如演化论）是合理的，那么我们中的大多数人也许会认

* 本文系中央高校基本科研业务费专项资助项目（编号 20720181103 及 2072019041）的阶段性成果。

① 两点说明：1. 尽管当代关于他心问题的讨论对象涉及的主要是他人，但是严格说来，这只是关于他心问题的狭义理解。广义的他心问题所涉及对象不应当仅限于除我之外的其他人（即“他人”），还应当包括其他非人对象（即“他物”）。本文的探讨适用于广义的他心问题。在文中，笔者将用“他者”（既包括他人也包括他物）来表示他心问题所涉及的全部对象。2. 笔者用“心灵状态”或“意识状况”意指各种常见的（有意识的）心理现象。按照目前学界的通行理解，它们既包括意向性的也包括现象性的。

为,某些类人物种(如大猩猩、黑猩猩等高等哺乳类动物),以及某些在演化树上的位置比类人物种距离人类稍远的物种(如猫、狗、海豚等哺乳类动物,或者鹦鹉、乌鸦等鸟类动物),不仅具有心灵活动,而且与人类心灵活动应当部分地相类似。但是,对于那些在演化树上的位置离人类相差甚远的物种(如草履虫等单细胞原生生物,或者蚯蚓等环节动物,或者蜻蜓等昆虫),我们一般不会觉得它们有心灵活动。或者,即使觉得它们中某些具有复杂行为的生物(如章鱼等软体动物)似乎具有心灵活动,那也只是一种与人类心灵相距甚远的心灵类型。而对于那些人造的物种(如未来可能出现的高智能机器),以及生活在遥远星系且与人类物理基质不同的地外生物(如可能存在的半人马座硅基生物),大多数人也许会认为,它们要么不具有任何心灵活动,要么即便有,也与人类的心灵类型相距甚远。

以上描述的是大多数人对他的常识看法。他心的怀疑论者恰恰要质疑这种常识,其质疑的立足点则基于“我心与他心在认知上有别”这样一个基本信念:

信念1: 在认知上,了解我自己心灵状态与了解他者心灵状态之间存在一个十分显著的区别: 我自己的心灵状态是可直接获知的,与之不同的是,他者的心灵状态则不是可直接获知的。

看上去,信念1能为大多数人所接受。在众多相关文献中,上述区别也常被表述为,认知者对自己心灵状态具有一种认知上的优先性或权威性。

然而,对某些极端的他心怀疑论者而言,如果信念1是合理的,那么如下情形就是可能的: 除我之外,其他个体完全没有任何心灵状态或意识活动,即我是世间唯一一个具有心灵的个体。换句话说,在他心的极端怀疑论者看来,如果我们排除不了上述这种可能,那么关于问题1的常识看法就并未得到有效辩护。

事实上,鲜有人会如此极端,关于他心的极端怀疑论毕竟离我们的常识太遥远。可以说,大多数他心的怀疑论者还是比较“温和的”。这些温和怀疑论者往往不质疑他人有心灵,也不质疑某些非人对象有心灵,他们要质疑的是,如下这两个命题得不到有效辩护:

命题1: 除我之外,其他人具有或能够具有与我类似的心灵状态。

命题2: 除了人类之外,某些非人对象的心灵与人类心灵之间具有或能够具有某种程度上的类似性。

命题1表述的是一个近乎常识的信念。从文献上看,当前绝大多数关于他心问题的争论基本上也是围绕命题1而展开的,但这并不意味着命题2不值得我们关注。若仅仅考虑命题1,我们得到的充其量只是他心问题的狭义理解。若要获得他心问题的广义理解,就需要将命题2也纳入到考虑的范围之内。

关于命题1的怀疑论,学界常见的两大应对方案分别是推论主义(inferentialism)和非推论主义(non-inferentialism)。以下两节先论述无论是推论主义还是非推论主义,目前都未能令人满意地应对关于命题1的怀疑论。然后,第四节前半部分将提出一个新方案,文中称之为“复合方案”(Hybrid Solution)。如果复合方案成立,那么命题1是有望得到有效辩护的。第四节后半部分会简要考察有关命题2的怀疑论,并论证复合方案对于命题2的辩护依然有效。最后一节将加以总结并给出最终结论。

二、应对怀疑论的进路之一: 推论主义

在推论主义的阵营中,如下两类方案最为常见: 类比推论方案(Analogical Inference Solution,

AIS), 以及最佳解释推理方案 (Inference to the Best Explanation Solution, IBES)。

AIS 的推理模式如下:

- (1) 对我而言, 我具有的心理活动 $M1$ 导致了我做出行为 $B1$ 。($M1 \rightarrow B1$)
- (2) 可以观察到, 他人 S 也做出了类似的行为 $B1$ 。($B1$)
- (3) 由 (1) 和 (2), 可以推出 (infer to) S 也具有类似的心理活动 $M1$ 。(($M1 \rightarrow B1$) $\wedge B1 \Rightarrow M1$)

对 AIS 有两个常见质疑。第一个是“逻辑不可检验性质疑” (logical uncheckability objection)。大意如下: 在不少学者看来, 上述推理明显不是演绎有效的。因为, 结论 (3) 中“推出”一词的意思 (即符号“ \Rightarrow ”的含义) 不是衍推, 而是类比, 属于归纳推论, 因而结论 (3) 的推理过程不具有逻辑保真性。在温和怀疑论者看来, 如果得到结论 (3) 的推理过程不具逻辑上的保真性, 那么 AIS 的支持者就需要提供逻辑之外的其他依据 (如某些经验上的可靠证据) 来表明这个推理过程仍然是有效的, 否则 AIS 就不能算做是一个关于命题 1 的有效辩护。“其他依据”如何获得呢? 一个自然的回答就是: 针对他人 S 的心理活动或心灵状态 $M1$ 再做经验科学上的检验。然而, 如果 AIS 的支持者试图对“ S 是否具有 $M1$ ”这一点做经验上的检验, 一个真正的麻烦就会出现: 由于他人 S 的心灵状态是逻辑上不可检验的, 因而试图从经验科学上来检验“ S 是否具有 $M1$ ”则是完全不可能的。对于观察者而言, 由于他人 S 的心灵状态不是客观的, 而检验或检测却是公共的或客观的手段, 因此检验 S 的心灵状态在逻辑上是不可能的。第二个是“孤例质疑” (one-case objection)。大意如下: 我观察到的从来只是自己的心灵状态 $M1$ 以及特定的行为 $B1$ 。一方面, 对我自己而言, $M1$ 和 $B1$ 这二者的确恒常地相继出现。但是另一方面, 我却从未曾观察到或检测到如下情况: 对他人 S 而言, $M1$ 和 $B1$ 这二者也恒常地相继出现。而且, 又由于 S 的心灵状态逻辑上不可检测, 因此, 我关于 $M1$ 与 $B1$ 之间关系的“研究样本”似乎永远只有我自己这一个孤例。显然, 仅凭一个孤例不可以归纳出“ $M1$ 导致 $B1$ ”这样一条 (普遍性的) 规律或似律陈述。正如维特根斯坦所言, 在他心问题上“仅由我自身情况这一孤例而类推他人”这种类比推论的思路 (AIS) 不但是无效的, 而且试图将此孤例普遍化的做法也是“不负责任的”。(维特根斯坦, 第 106 – 107 页)

上述两个质疑的确击中了 AIS 的要害。但推论主义的支持者认为, 上述质疑也许揭示了 AIS 存在缺陷, 但不能由此断定, 推论主义的基本立场是有问题的。近年来, 如下两个推论主义新方案备受关注, 一个是最佳解释推理方案 (IBES), 另一个则是诉诸自然齐一性 (Uniformity of Nature) 的类比推论方案改进版 (UN-AIS)。

IBES 的思路如下: 在承认“逻辑上的不可检验性”和“孤例类推无效”的前提下, 对后者做进一步分析。首先, 区分出两种不同的孤例推论模式: 一类是孤例类推推论, 另一类则是孤例科学推论。其次指出, 对于他心的推论不是一种类推推论的模式, 而是一种特殊的科学推论的模式, 即最佳解释推理模式。因此, 虽然对于他人心灵状态的推论是孤例推论, 但却不是孤例类推, 而是一种基于孤例的最佳解释推理 (one-case based IBE)。因此, 对“他人 S 具有行为 $B1$ ”的最佳解释就是: $B1$ 是由一个不可观察的理论实体 $M1$ 导致的。(cf. Pargetter) 换句话说, 对观察者 O 而言, 尽管没有逻辑上的必然性来保证, S 做 $B1$ 是由 S 所具有的心理活动 $M1$ 所导致的 (因为 $M1$ 有可能并不是 $B1$ 的唯一的充分条件), O 也不可能通过归纳推理得出这一点 (因为孤例类推无效), 然而 O 仍有理由相信结论 (3) 为真。理由就是, 得出 (3) 的推理过程 (\Rightarrow) 是一种科学的推论过程, 即最佳解释推理。因此, 对于 S 做出 $B1$ 的最佳解释就是, S 具有 $M1$ 。

以下这两个问题似乎是 IBES “绕不开的”：第一，M1 的因果效力问题。IBES 成立的一个前提是，M1 有资格成为 B1 的充足原因。如果这个必要前提不成立，那么 IBES 就失效了。换句话说，作为一种心灵状态或心理活动，如果 M1 不具有真正的因果效力，那么 M1 就不可能成为 B1 的充足原因（因为 B1 的充足原因有可能是物理的，而非心理的）。那么，M1 是否具有真正的因果效力呢？这个问题涉及到心灵哲学里的一个难题——排他性问题（exclusion problem），目前未有定论。因此如果 IBES 成功与否最终要依赖于 M1 是否具有真正的因果效力，那么 IBES 就需要面对排他问题。第二，最佳解释的选取标准问题。大意如下：一方面，如果我们愿意接受认知神经科学（及其相关经验研究）成果，那么就有理由相信，M1 总需要有某个神经回路 N1 来作为其出现的物理基质，M1 才能出现。换句话说，一方面，神经科学告诉我们，不管是观察者 O 还是他人 S，只要 N1 出现，M1 就会出现，另一方面，按照 IBES 的思路，S 做出 B1 是由于 S 具有 M1。那么，现在问题来了：在观察者 O 看来，对于“S 做出 B1”的最佳解释，究竟是 M1，还是 N1？

IBES 的支持者或许会说，出于交流便利性或解释简单性的考虑，M1 是“最佳的”。但对一个认知神经科学家而言，首要考虑的可能是理论的真实性和非交流的便利性。如果实际情况是 M1 并不具有真正的因果效力，导致 B1 出现的原因乃是 N1，那么把 M1 作为 B1 出现的原因，就不仅不是最佳的，很可能还是错误的。出于理论真实性的考虑，我们似乎更应相信 N1 才是“S 做出 B1”的最佳解释。可见，实际情况并非如 IBES 的支持者想的那么简单。事实上难以找到合适的标准来断定究竟是 M1 还是 N1，才是“S 做出 B1”的最佳解释。正如希斯洛普（Alec Hyslop）所言，除非可以排除掉 N1 是 B1 的原因，才能一并排除掉“N1 是‘S 做出 B1’的最佳解释”这一可能性。（cf. Hyslop, 1995, p. 37）此外，IBES 至少还有如下两种情形需加以排除：查莫斯（David. J. Chalmers）的僵尸（Zombie）（cf. Chalmers），以及休梅克（Sydney Shoemaker）的颠倒光谱（inverted spectrum）（cf. Shoemaker）。如果 M1 特指感受质，那么在僵尸情形中，“僵尸 S2 做出行为 B1”的最佳解释可能是 N1，而非 M1。而在颠倒光谱情形中，“颠倒光谱症患者 S3 做出行为 B1”的最佳解释则可能是 M2，而非 M1（M2 是不同于 M1 的现象心灵状态）。在 M1 的因果效力问题，以及最佳标准选取问题都没得到合理解释之前，我们似乎没有理由相信，IBES 算是一个（关于命题 1 的）有效辩护。

下面看 UN-AIS，其思路如下：与 IBES 一样，UN-AIS 也承认上述两个质疑（“逻辑上的不可检验性”和“孤例类推无效”）的合理性，同样也认为，对于他人心灵状态的推论是孤例推论。二者的重要区别则在于：IBES 认为这里的孤例推论不是类比推论，而是一种科学的推论，是基于孤例的最佳解释推理，而 UN-AIS 则认为这里的孤例推论仍是类比推论，“实际上，这是诉诸自然齐一性”的类比推论，因而“不是类似的结果有类似的原因，而是类似的原因有类似的结果”。（Hyslop, 1995, p. 53）

对于 UN-AIS 的支持者而言，正是自然齐一性原则保证了结论（3）为真。换句话说，由（1）和（2）能推出（“ \Rightarrow ”）结论（3），这里“推出”一词的意思（即符号“ \Rightarrow ”的含义），既不是最佳解释推理，也不是某种逻辑推理，而是运用了自然齐一性原则。因此这仍是个有效的类比推论。可见 UN-AIS 似乎不但守得住类比推论的立场，而且能避开最佳标准问题的困扰。因此希斯洛普乐观地认为，诉诸自然齐一性（UN-AIS）乃是“推论主义值得我们期待的唯一方式”。（Hyslop, 2014, § 3.1）

对此，笔者有两个疑虑。第一，如果自然齐一性原则说的是“同因则同果”，那么在此运用该原则就有一个必要的前提条件：仅当 M1 和 B1 之间具有因果关系。换句话说，除非 M1 具有真正的因果效力，否则自然齐一性原则在此就不适用。可见，UN-AIS 同样要面对 M1 是否具有因果效力的问题。第二，我们并不清楚自然齐一性原则是否也适用于心灵。如果自然齐一性原则也适用于心灵，那么就会有心理-物理规律或心理-心理规律。然而科学哲学和心灵哲学里有大量论述表明，很可能并

不存在心物规律和心心规律。例如，刚刚提到的僵尸论证和颠倒光谱论证，就是两个著名例子。而如果不存在心物规律和心心规律，那么自然齐一性原则也就不适用于心灵。可见，UN-AIS 还要面对心物规律和心心规律是否存在的问题。因此，在上述两个疑虑得到令人满意的解答之前，UN-AIS 还不算是个（关于命题 1 的）有效辩护。

三、应对怀疑论的进路之二：非推论主义

在非推论主义方案中，如下两类较为常见：一个是标准方案（*criterial solution*，CS），另一个是态度方案（*attitudinal solution*，AS）。

CS 有两个要点：第一个要点是认为，“一个‘内在的过程’需要外在的标准”（维特根斯坦，第 166 页），即“行为应当被看做是心灵状态在场的标准”（Hyslop, 2017, Section 3.3）。从认识论层面讲，行为与心灵状态之间不是归纳的推论关系，而是具有一种“非推论的联系或概念上的连接”（*ibid.*）。仍以上述情形为例。只要出现特定的行为 B1，就标志着行为者具有特定的心灵状态 M1。因此在 CS 的支持者看来，推论主义者的错误在于“他没能把行为当做是心理现象的标准”。只要看清这一点，就会明白，在他心问题上，不管是类比推论（AIS 和 UN-AIS）还是科学推论（IBES）都是不适用的。（*cf.* Malcolm, p. 975）

CS 的第二个要点是认为，对自我心灵状态的认知不具有认知上的优先性。从上面第一点可推出第二点。由于行为是心灵状态的标准，而标准具有公共可观察或可检测的特征，因此，不管行为者是谁，认知其心灵状态的方式都是一样的。因而命题“我有心灵状态 M1”与命题“他人 S 有心灵状态 M1”，所获辩护的依据或理由应当是同一类。这是个相当惊人的结论。如果它是合理的，那么本文第一节所提到的信念 1 就是虚假的。由于信念 1 是他心怀疑论据以攻击我们关于他心的常识看法的“立足点”，如果能够成功论证信念 1 是虚假的，那么这将是一个“釜底抽薪式”的论证，他心的怀疑论将因此而被“消解掉”。^①

CS 果真能以如此彻底的方式消解掉他心的怀疑论么？学界对此存有争议。最为常见的有如下三类质疑：标准太弱、CS 冗余，以及行为主义翻版。先看标准太弱。很多行为似乎不足以充当心灵状态的标准或证据。例如，闭目静思这个行为。即使观察到他人 S 闭目静思，一个观察者似乎仍不足以知道 S 此刻的内心状态究竟如何。类似的例子在生活中比比皆是。CS 似乎没法令人满意地解释这些反例。正如希斯洛普所言，“如果既不存在任何归纳的推论关系，也不存在逻辑的衍推关系，那么可观察的行为与不可观察内在状态之间的那条鸿沟，是不可能仅仅通过下命令的方式而被跨越过去的”。（Hyslop, 2014, § 3.2）

CS 的支持者或许认为，如果行为 B2 不足以充当心灵状态 M1 的标准，那么，错的不是 CS，而是观察者没有找准行为。这样似乎就能回应关于标准太弱的质疑。由于行为 B1 与 M1 之间的概念联系不是偶然的或任意的，因此一旦找准了行为（即 B1），那么观察者就应当能“仅仅在行为中看到心灵状态”（Spaulding, p. 472）。这个回应会招致如下两个质疑：CS 冗余，以及 CS 乃是行为主义的一个“翻版”。

下面看 CS 冗余。如果我们接着问：如何才算找准了 B1？CS 的支持者会说，只要在恰当的情境

① 近年来，在认知科学和知觉哲学领域，他心的直接社会感知理论（*direct social perception*，DSP）备受关注。正如斯波尔丁（S. Spaulding）所概括的，“DSP 基本想法是，我们能够以一种非推论的方式察觉到他人的心灵状态”，因而“依据 DSP 理论，我切实地察觉到了特定的忧愁、恐惧和欲求。我并不是先观察到了特定的行为然后推出特定的心灵状态，而是仅仅在行为中就看到了心灵状态”。（Spaulding, p. 472）可见，DSP 的基本思路类似于 CS。本文对 CS 的分析同样也适用于 DSP。

中,即观察到 B1 发生时的满足条件 C1,我们就可以依据 C1 来断定 B1 出现了。但是,若我们继续追问“什么样的情境才算恰当的,或者为何 B1 的满足条件是 C1 而非其他”这类问题,就会发现,如要解释清楚“恰当的”一词的准确意思,或解释清楚为何 C1 是 B1 出现的满足条件,就不得不再诉诸新的依据 C2 来解释“何为恰当的情境”,或进一步解释 C1 的合理性。可是,如果我们再继续追问 C2 的合理性来源,那么 CS 的支持者可能有如下两个选择: 1. 要么继续给出另外的依据 C3。然而,若沿此思路,我们一直追问下去,CS 的支持者就需要不断地寻求新的依据,而这会导致无穷倒退; 2. 要么诉诸某种约定 T。对于观察者 O 而言,T 可以是外部的也可以是内部的。前者可以是某个(科学共同体给出的)定义,也可以是某种历史-文化-社会习俗,而后者则可以是 O 所具有的某种心灵状态或心理意象。由于 T 要么是约定的,要么是类似于公理的定义,要么是 O 自身的某种心灵状态或心理意象,因此一旦诉诸约定,我们就不能再继续追问 T 的合理解释了(在此 T 成为了一切解释的最终依据,是解释的“尽头”)。可见,诉诸 T,CS 的支持者似乎有效地阻止了无限倒退。然而,诉诸 T 依然挽救不了 CS。理由就是:不难发现,对 O 而言,不管 T 是外部的还是内部的,T 与 B1 之间总要具有一种认识上的恒常联系。也就是说,如果 T 出现,那么 B1 出现,并且,如果 T 不出现,那么 B1 也不出现($T \leftrightarrow B1$)。对于 CS 的支持者而言,正是由于这二者之间具有这种认识论上的恒常联系,我们才能把 T 当作是 B1 出现的恰当情境或满足条件,或者用 T 来定义 B1。换句话说,对 O 而言,只要 T 一出现,那么 O 就可以断定行为 B1 出现了。现在问题来了,“对 O 而言,只要 T 一出现”是什么意思?或者说,O 是如何察觉到 T 出现的?答案似乎有且仅有如下两个: i. 如果 T 是外部的,那么 O 则要诉诸证实主义的基本方法或手段: 要么凭借某种物理的或科学的验证程序,将呈现在自己面前的情境记录或拍摄下来,并与 T 的物理样本 T^* 进行物理比对,如果发觉二者是(在误差允许的范围内)一样的,那么就证实 T 出现了。要么 O 将观察所获得的感觉材料与自己以往记忆中关于 T 的心理意象或心理样本 T^{**} 进行比对,如果发觉二者是一样的,那么也证实 T 出现了。 ii. 如果 T 是内部的,即 T 本身就是 O 的某种特定的心灵状态或“心理样本” T^{**} (即 $T = T^{**}$),那么 O 只需将观察所获得的感觉材料直接与 T^{**} 进行比对,如果发觉二者是一样的,那么 O 就证实 T 出现了。不难看出,这里所发生的认知过程与上述证实主义的第二种情形是一样的。可见对观察者 O 而言,要察觉到 B1 出现的恰当情境 T,要么诉诸某种物理的比对程序,要么诉诸某种心理的比对程序。“O 察觉到 T 出现”的认知过程或推理模式如下:

- (1') 对观察者 O 而言,如果 T 出现,那么 T^* 或者 T^{**} 。($T \rightarrow (T^* \vee T^{**})$)
- (2') 依据证实主义的手段, O (将自己获得物理数据或者感觉材料与特定的物理样本或心理样本进行物理的或心理的比对之后) 证实, T^* 或者 T^{**} 出现了。($T^* \vee T^{**}$)
- (3') 由 (1') 和 (2'), 可以推出 T 出现了。($(T \rightarrow (T^* \vee T^{**})) \wedge (T^* \vee T^{**}) (\Rightarrow T)$)

不难看出,上述推理模式与推论主义的 AIS 模式是一样的,都是归纳的类比推论。可见,CS 诉诸约定 T 是无济于事的。换句话说,观察者 O 诉诸 T 的一个必要条件是, O 能够察觉到 T 出现,然而 O 察觉 T 出现的认知模式却是推论主义的(归纳的类比推论),因此诉诸 T 的最后结果就是,我们竟然背离了非推论主义的基本立场而走上了推论主义的进路,而这也意味着,将 CS 称作是非推论主义的方案实属多余。

第三个质疑认为,CS 其实就是行为主义的一个翻版。如果行为与心灵之间的联系不是偶然的,而且观察者能“仅仅在行为中看到心灵状态”(Spaulding, p. 472),那么一个自然的理解是,心灵状态其实就是行为。这是行为主义的基本观点。因而有学者认为,CS 其实是一个“粗糙的行为主义”

(crude behaviorism) (Jacob, p. 531)。对此批评,一个可能回应是,首先指出,心灵与行为之间的关系即使不是偶然的,但也不一定就是等同关系或构成关系,因此不能做行为主义的解释。而是应当理解成是一种弱的构成关系或部分的构成关系。正如冰山尖与冰山的关系。冰山尖是冰山的构成部分,但前者显然不等同于后者,或者后者也不是完全由前者构成。(cf. Krueger, p. 8) 可见 CS 与行为主义并不相同。依据这里的解释,CS 的推理模式如下:

- (1'') 如果行为 B1 出现,那么他人(即该行为的做出者) S 具有心理活动 M1。(B1→M1)
- (2'') 观察者 O 察觉到, S 做出了行为 B1。(B1)
- (3'') 由(1'')和(2''),可推出(entail)他人 S 具有心理活动 M1。((B1→M1) ∧ B1→M1)

上述推理模式的确有别于行为主义,关键区别就在于第一步((1''))。从逻辑上看,行为主义会认为 B1 与 M1 之间是一种双向衍推关系(\leftrightarrow),而依据上述解释,B1 与 M1 之间的关系,则应当被理解成单向的衍推关系(\rightarrow)。然而这个解释又会使得 CS 陷入“冗余论”的质疑,关键在第二步((2''))。这一步说的是“O 察觉到 S 做出了行为 B1”。不难理解,对于 CS 的支持者而言,这个认知过程或推理模式也一样要诉诸证实主义的基本方法,而这与上述刚刚给出的“O 觉察到 T 出现”的推理模式如出一辙。^①

可见,只要把行为看做心灵的标准,上述察觉过程(无论是觉察到 T,还是觉察到 B1)就是 CS 得以成立的一个必要前提,^②因此,即使能避开“标准太弱”的质疑和行为主义的质疑,CS 仍避不开“冗余论”的质疑。

这的确击中了 CS 要害,非推论主义者们转而寻求替代方案,态度方案(AS)就是这样一种替代方案。AS 认为,“他人有心”不是一个信念,而是一个不同于信念且独立于信念的态度。这种态度不仅在心理学上而且在认知上比信念“更深”(Hyslop, 1995, pp. 124 – 125)。正如希斯洛普所言,AS 的基本主张是“我们并不相信‘他人有心’。我们有的只是一个关于他心的态度,即他人被看做与我类似,他人的体验亦被当做与我的体验类似而已”。(ibid., p. 125) 因此,之所以观察者 O 认为,他人 S 做出 B1 时具有与自己在做出 B1 时一样的心灵活动 M1,不是因为 O (按照推论主义的模式)从 B1 推知 M1,也不是因为 M1 与 B1 之间具有概念上的非偶然联系或部分构成关系,使得 O 可以把 B1 当作标准,从而仅仅在 B1 中看出 M1 (像 CS 所宣称的那样),而是因为, O 具有一种非信念的态度,从而能直接把某个与 O 身形类似的对象看做是一个自己的同类(即看做是 S),并且,还能直接地把 S 的某些与 O 类似的举动看做是由 M1 所导致的 B1。这里, M1 也是导致 O 自身做出 B1 的充足原因。

在 AS 的支持者看来,怀疑论的错误在于没有认识到我们“具有这样一种对待事与人的态度,这种态度比推论性的信念更深更直接”。(Hyslop, 2014, § 3.3) 由于这种非信念的态度不是认识论层面的,因此怀疑论者不应当要求我们从认识论层面给出关于他心的辩护。我们从未具有“关于他心的信念”,我们一直具有的是“关于他人以及他心的态度”。这种非信念的态度一直“如其所是地内嵌在(built in)我们与其他个体的交往之中”。(ibid.) 一旦明白了这一点,怀疑论就会因失去了其质疑的“立足点”而立刻消失。

① “O 觉察到行为 B1”的推理模式可以概括如下: $(B1 \rightarrow (B1^* \vee B1^{**})) \wedge (B1^* \vee B1^{**}) \Rightarrow B1$ 。其中, $B1^*$ 与 $B1^{**}$ 依次指 B1 的物理样本和心理样本。不难看出,这也是一个与 AIS 推论模式一样的归纳的类比推论。

② 无论是推论主义方案还是非推论主义方案,观察者察觉他人行为(或行为出现的恰当情境)的认知过程都是其中一个必不可少的前提。虽然说,对此认知过程的理解可以存在差异,但是这个认知过程应当成为所有关于他心问题的方案的一个共同的必要前提。希斯洛普也有类似表述:对任何一个关于他心问题的理论而言,观察者或“理论者本人的经验都是关键的”。(cf. Hyslop, 2014, § 4)

凭上述消解策略, AS 似乎避开了怀疑论的质疑, 但如下两个问题难以回避: 第一个问题是, 他心态度的可错性问题, 即我们关于他心的态度是否可能出错。希斯洛普指出, 某些非认识论的态度似乎是错误的。例如, 种族歧视或者性别歧视。(cf. Hyslop, 2014, § 3.3) 如果关于他心的态度不同于种族歧视或性别歧视, 是一种不会出错的普适态度, 那么似乎需要对“为何我们关于他心的态度可以如此特别”提供进一步的解释。否则 AS 会被认为是特设的 (*ad hoc*)。对此, AS 的支持者很可能会援引维特根斯坦的生活形式等有关论述来做进一步解释: 之所以我们直觉上会觉得性别歧视或种族歧视主义者秉持的是错误的态度, 是因为我们与他们分别处在不同的生活形式之中。因此, 这里不是由同一种生活形式内部的分歧所致的错误, 而是由不同生活形式所致的差异。然而这个回应会招来更大的麻烦, 即第二个问题。AS 的支持者面临的第二个问题是, 这个回应消除不了如下这个直觉上的困惑: 一方面, 我们似乎会认同, 亚马逊雨林深处的原始部落与我们在对待具体行为举止的态度上的明显差异, 是由于各自生活形式明显不同所致的差异。另一方面, 我们似乎不会认同, 种族歧视论者和性别歧视论者与我们在对待具体行为举止的态度上的明显差异, 也是由不同的生活形式所致。我们大多认为, 种族主义和性别歧视是完全错误的, 而非我们与他们(在对待人与物的态度上)有所差异。事实上, 很多种族主义者和性别歧视论者不仅和我们共享许多基本的常识, 他们中的一些人甚至还曾经是我们中的一份子并且和我们受过一样的良好教育。因此很难想象, 他们与我们分属于不同的生活形式。换句话说, 这里有个直觉上的困惑: 诉诸生活形式之间的差异, 似乎可以满意地解释亚马逊人和我们之间的分歧, 而同样诉诸生活形式之间的差异, 为什么我们会觉得, AS 解释不了种族主义者或性别歧视论者与我们之间的分歧呢?

这个困惑似乎并不能仅凭诉诸生活形式就完全消除。可以预料, AS 将面临如下这个两难: 一方面, 如果诉诸生活形式并不能消除掉上述困惑, 那么仍认为这个困惑是生活形式的一部分因而不能获得进一步解释, 就会使 AS 陷入神秘主义或不可知论; 另一方面, 如果认为上述困惑有望获得进一步解释, 那么我们就可以对此解释的合理性继续进行追问, 而这又会使得 AS 面临 CS 类似的结局: 要么陷入无限倒退, 要么变成冗余。

AS 的支持者与我们或许最终能达成如下妥协: 一方面, 他们不得不承认, 在“他人具有何种特定的心灵状态”这一点上总是有出错的可能。另一方面, 我们也得承认, 在“他人是具有心灵的”这一点上的确没有出错的可能。因为, 我们有“关于他人具有心灵”的非认知态度, 却没有“关于他人具有怎样的心灵状态”的非认知态度。必须承认, AS 的确给我们带来了有益启示, 但同时也应看到, AS 仍不算是一个(关于命题 1 的)有效辩护。

四、处理他心问题的复合方案

以上分析表明, 无论是推论主义还是非推论主义, 目前看来, 似乎都难以给出关于命题 1 的有效辩护。如果我们仍试图寻找关于命题 1 的有效辩护, 那么可能的选项大概仅有如下三个: a. 继续坚持推论主义进路; b. 继续坚持非推论主义进路; c. 寻找其他方案。

该做什么选? 笔者目前的看法是, 选择 c 是最为可取的。理由如下: 一方面, 推论主义除了要回应所面临的那几个质疑(见本文第二节)之外, 有个关键问题是推论主义原则上无法处理的。那就是, 所有的推论主义方案都有一个共同的前提: 他人是有心灵活动的。换句话说, 推论主义关心的是, 心灵活动与行为之间是怎样的推论关系。但是, 任何一个关于心灵状态与行为之间的推论主义方案都必须以“他人是有心灵的”这一点作为其必要的前提。问题是, 该前提在推论主义的框架下得不到辩护。另一方面, 非推论主义对“他人是有心灵的”这一点能给出合理解释。AS 给出的解释是, “他

人具有心灵”这一点是我们所具有的一种非认知态度，而不是一个认识论上的信念。但是上述分析显示，非推论主义对于特定行为与特定心灵状态之间具体联系的非推论主义说明却是不成功的。具体来说，CS 即使可以避开标准太弱质疑，也可以避开行为主义质疑，却避不开冗余论质疑。而 AS 最终也不得不承认，即使诉诸非认知态度，也无法保证我们在“他人具有何种心灵状态”这一点从不犯错。因此，推论主义一方无法为其前提“他人具有心灵”提供辩护，却有望为“特定行为与特定心灵之间具有何种关系”提供辩护。而非推论主义一方则可以为“他人具有心灵”提供辩护，却难以以为“特定行为与特定心灵之间具有何种关系”提供辩护。可见，无论是推论主义，还是非推论主义，任何一方单独地都不足以给出关于命题 1 的有效辩护。或许以某种方式把推论主义和非推论主义各自的优势相结合，才有望找到一个关于命题 1 的有效辩护。因而，相对于 a 或 b，选择 c 是最为可取的。选择 c 意味着，我们打算采取的辩护进路既不是推论主义的，也不是非推论主义的，而是两者的一个复合。这条新进路下的具体方案可能不止一个，但都可以统称之为“复合方案”。下面笔者先给出一个具体的复合方案，然后再与推论主义和非推论主义的方案进行比较，看看复合方案可能具有哪些优势。该复合方案的推理模式如下：

- (1*) 他人 S 具有心灵。(关于他心的态度)
- (2*) 心理状态 M1 导致观察者 O 做出行为 B1。(M1→B1)
- (3*) 对于一个具有心灵的个体而言，如果其行为 B1 有一个心理上的或经验上的充足原因 M1，那么 B1 也会有一个认识论上的充足原因 P1，并且 P1 = M1。(token-identity principle)
- (4*) O 察觉到 S 做出了类似的行为 B1。(B1)
- (5*) 类似的原因总是导致类似的结果。(uniformity of nature principle)
- (6*) 由 (3*) (4*) (5*) 可得，类似的原因 P1 导致 S 做出 B1。((P1→B1) ∧ B1 ⇔ P1)
- (7*) 由 (1*) (3*) (6*) 可得，类似的原因 M1 导致 S 做出 B1。((P1 = M1) ∧ (P1 ⇔ B1) → (M1 ⇔ B1))
- (8*) 由 (2*) (7*) 可得，S 具有(与 O 在做出 B1 时)类似的心理活动 M1。

下面分析该论证过程。前提 (1*) 是 AS 带给我们的有益启示。用维特根斯坦的话来说就是，“我对他的态度是对心灵的态度，并非我认为他有灵魂”。(维特根斯坦，第 195 页) 可见，(1*) 是我们具有的一种关于他心的非认识论的态度，也可以理解为一个属于观察者 O 和他人 S 共同生活形式的语法命题。同样地，前提 (2*) 也可看做是一个语法命题。(2*) 的确定性基于信念 1，即 O 能直接获知自己做出 B1 的原因是 M1。前提 (3*) 的意思是，如果一个有心灵的个体宣称，从心理上或经验上找到了自己做出 B1 的动机或原因 M1，那么，O 就能从认识论上找到一个关于 B1 的原因 P1，且使得 P1 = M1。也就是说，B1 的出现如果有一个心理上的动机或充足原因，那么，相应地就会有一个认识论上的充足原因。

注意：(3*) 包含两个要点：第一个要点是，对于任何一个导致 B1 出现的事件 E，可有两种不同的言说方式：一种是心理上或经验上的，一种是认识论上的。因此，如果采取心理上的言说方式来解释为何 B1 出现，就会把 E 说成是一个该行为者主观上具有的心理动机或心理原因 M1。如果采取认识论上的言说方式来解释为何 B1 出现，那么就会把 E 说成是一个客观上可被测量的或可获得公共理解的认识论原因 P1。M1 和 P1 只是关于同一个 E 的两种不同说法而已。可见，M1 = P1 之间的符号“=”的意思是“共指称的”，即用心理描述和用认识论的描述说的是同一回事 (E)。第二要点是，

M1 和 P1 之间的这种共指称关系，不是逻辑上必然的，也不是形而上必然的，而是律则上必然的。^①因此，尽管 M1 与 P1 是共指称的，但任何一条关于 B1 出现原因的心理描述都不能被还原成关于 B1 出现原因的认识论描述。换句话说，当且仅当严格的或无例外的桥接律或者对应规则存在，这种还原或翻译才是可能的，然而严格的或无例外的桥接律或对应规则只存在于认识论描述之间，因此在关于 B1 出现原因的心理描述和认识论描述之间不存在这样的桥接律或者对应规则。

结合以上两点，可以得出，作为 B1 的一个心理原因 M1，会存在一个相应的认识论原因 P1，且 $P1 = M1$ ，但是 B1 的所有心理原因所构成的类（type）并不等同于 B1 的所有认识论原因构成的类。可见，（3*）表达的是一种殊型—殊型间的同一（token-token identity），而非类型—类型间的同一（type-type identity），因而可以把（3*）称为“殊型同一原则”。

下面看前提（4*）。无论是推论主义者还是非推论主义者应当都能接受该前提。争议之处在于，O 的察觉模式究竟是怎样的，特别是，其中是否有一个比对过程发生。推论主义者会认为有，因而察觉模式是类推的。而非推论主义者会反对，因而察觉模式不是推论的。由于无论哪一方有道理，都不影响复合方案的结论，因而我们可把（4*）作为复合方案的一个前提。

前提（5*）是自然齐一性原则。尽管也存在关于该原则究竟是归纳的还是非归纳的争论，但同样也不影响复合方案的结论，因而我们可把（5*）也作为复合方案的一个前提。

（6*）运用了自然齐一性原则。由于 O 察觉到 S 做出类似的行为 B1，依据自然齐一性原则，类似的原因 P1 导致 S 做出类似的行为 B1。注意一点：由于 P1 和 B1 都是客观可测的或可公共观察的，因而自然齐一性原则在此完全适用。

下面看（7*）。由于 O 和 S 做出类似行为 B1 的认识论上的原因是类似的（（6*）），又由于 S 是一个有心灵的个体（（1*）），因此殊型同一原则对 S 适用。因此，按照殊型同一原则，O 和 S 做出类似行为 B1 的心理上的原因也应当是类似的（（3*））。这里要注意两点：I.（7*）运用了殊型同一原则，即 M1 和 P1 之间具有一种律则上的必然关系。换句话说，按照这种律则必然性，如果两个有心灵的个体（O 和 S）做出行为 B1 的认识论上的原因是类似的（P1），那么对于 O 和 S 而言，行为 B1 同样也都具有心理上的原因，而且（二者做出 B1 的）心理上的原因也是类似的（M1）。不难看出，运用殊型同一原则，可以使复合论证成功避开“自然齐一性原则是否（原则上）适用于心灵”的质疑。相对于 UN-AIS 而言，这是个明显的优势。II. 要得到（7*），前提（1*）也是不可缺少的，否则既无法运用殊型同一原则，也排除不了极端怀疑论。也正是因为缺少前提（1*），上述所有的推论主义方案才应付不了极端怀疑论。由于 O 在做出 B1 时的充足原因是 M1，而 M1 也是导致 S 做出 B1 的充足原因，因此由（2*）和（7*）就能得出，S 在做 B1 时与 O 在做出 B1 时具有类似的心理活动 M1，即结论（8*）。以上就是复合方案的整个论证过程。如果上述分析合理，那么我们应当有理由相信，复合方案关于命题 1 的辩护是有效的。

下面简要考察命题 2。命题 2 说的是，除人类之外，某些非人对象的心灵与人类心灵之间能够具有某种程度的类似性。对命题 2 也有两个不同版本的怀疑论：极端的和温和的。前者认为，非人对象具有心灵这一点得不到有效辩护。而后者则可以再进一步区分为两种不同类型。类型 1：我们无法确定“具有心灵的非人对象的范围”。类型 2：我们即使可以确定哪些非人对象具有心灵，也仍无法确

① 福多亦有类似看法。他认为“之所以（特殊科学）S 中的谓述与（基础物理学）P 中的谓述能用来描述同样的事物（即 S 中的谓述能描述的那个事物子集，P 中的谓述也同样可以），是因为这二者之间的关系（即 S 中的谓述与 P 中的谓述之间的关系）可以是律则上必然的。”（Fodor, p. 99）

定“具有心灵的非人对象与人类心灵的相似程度”。

不难看出,非推论主义方案能有效应对关于非人对象心灵的极端怀疑论。考虑如下问题:一个常年生活在深山老林中的护林人为何不怀疑与他朝夕相处的猎犬是有心灵的?凭借非信念态度所带来的启示,可以这样回答:护林人之所以不怀疑他的猎犬有心灵,是因为朝夕相处的时光使得他早已把这只猎犬当成了亲密伙伴,于是他就不再怀疑这个亲密伙伴是不是有心灵,哪怕有时候他也拿不准盯着炉火在发呆的这位“老伙计”究竟在琢磨什么。正是由于具有这种非信念态度,我们不但能够把那些在外形上或者物理上与我们相似的个体当作是自己的同类,我们似乎也能把那些在外形上和物理上与自己有所差异的非人对象当作是有心灵的。一旦这样做,我们就不会怀疑非人对象也是有心灵的。由此可见,关于他物(非人对象)心灵的极端怀疑论与关于他人心灵的极端怀疑论犯了同样的错误。

如果我们具有这种非信念的态度,那么似乎也能应对类型1。具体来说,如果我们只是凭借这种非信念态度,来把某些非人对象当作是有心灵的,那么,这同时就意味着,我们究竟把什么样的非人对象当作有心灵的,这一点并没有认识论上的依据。因此,对于类型1的回应是,如果我们的确具有非信念的态度,那么我们也就可以在经验上或者心理上区分出哪些非人对象可以具有心灵,哪些并没有心灵。凭借这种非信念的态度,非推论主义者能否再进一步,有效应对类型2呢?答案似乎是否定的。理由是,日常生活的经历告诉我们,我们中的大多数人,不但在揣测他人具体的心灵状态时都犯过错,而且在揣测非人对象的具体心灵状态的上,更是如此。可见,非推论主义的方案似乎难以应对类型2。另一方面,不难看出,推论主义方案似乎对他物心灵的极端怀疑论是无能为力的,对类型1也是无能为力的。然而,对于类型2,推论主义者似乎能有所作为。相对他人情形而言,尽管他物的情形也许更加复杂,但二者的认知过程本质上没有差别。^①可以料想,在殊型同一原则和自然齐一性原则的帮助下,对类型2给出某种推论主义方案的辩护也不是不可能的。可见,要给出关于命题2的有效辩护,很可能仍要采取复合方案。

① 两点说明: I. 之所以他物情形会异常复杂,主要是因为如下两个因素: 第一, 哪些非人对象是有心灵的; 第二, 有心灵的非人对象的哪些举动被当做与观察者O的哪些行为类似。对O而言, 关于第一个因素的确信, 来自于某种非信念的态度。关于第二个因素的确信, 则既可以来自某种非信念的态度为O所提供的精微莫测的(imponderable)经验性依据, 也可以来自认识论上可公共观测的证据。对于他人的情形, O往往只需要考虑第一个因素, 除非在某些陌生的或特殊的情境中, O一般不需要考虑第二个。而对于非人对象的情形, 上述两个因素, O都要考虑, 除非O对自己观察的非人对象已经相当熟悉(如, 上述护林人与他的猎犬的例子)。在考虑第二个因素时, O越是把非人对象Y的某些举止当做与O自己的行为B1类似, 那么O就越是会认为, Y具有(在O做出B1时)类似的心灵状态M1(如, 护林人与他的猎犬的例子)。反之, O要是觉得Y的举止怪异甚至难以看出Y的举动与O自己的哪些行为类似, 那么O就会认为Y的心灵状态与自己的心灵状态类似程度较低, 甚至认为Y的心灵是一种陌生类型(如, O是一个生活在格林兰岛上的爱斯基摩人, 当他第一次看到一只宠物绿鬣蜥时。或者, O是一个未来的地球人, 当他与来自遥远的半人马座的硅基智慧生命初次接触时)。但是, 也不能排除如下情形: 随着O与Y的朝夕相处或多次接触, O越来越熟悉Y的行为举止(即O越来越频繁地看出Y的哪些举止与自己的哪些行为是类似的), 从而也越来越了解Y的心灵。II. 这里有人或许会问: 按照上述这种认知模式, O是否最终能获得关于Y心灵的完整(complete)了解, 就像O了解他人S的心灵一样? 我的看法是, 从认知途径上说, 我们了解Y的心灵与了解他人S的心灵并没有本质的不同。因此, 如果我们觉得可以获得关于他人心灵的完整了解, 为什么就不能有关于Y心灵的完整了解呢? 这里的关键是, 如何理解“完整的”这个词。首先, 我们只能从认识论上来理解“完整的”这个词的意思。(因为, 正如维特根斯坦所言, 不可能有关于任一词项的私有理解。) 其次, 如果我们可以接受殊型同一原则, 并且信赖可能找到的认识论上的证据。那么有理由相信, 是可以做到完整了解的。只不过, 这种“完整了解”的意思是一种从认识论上所获得的完整理解。这里, 若有人觉得, 总有某种第一人称视角的感受质, 深藏在他者的内心中, 以至于观察者原则上无法从外来窥其究竟, 从而妨碍获得上述完整理解。(转下页注)

五、结论

至此，我们应当明白，广义的他心问题不仅涉及到了他人的心灵，也涉及到了那些非人对象的心灵，而且我们要应对来自他心怀疑论的挑战，就不能仅限于认识论层面的辩护，也要在认识论之外找寻必需的资源。

本文的结论如下：他心问题的确是个复杂的问题，任何单一的处理方案在它面前都难免捉襟见肘。也许该考虑换个思路了！要对付这样一个复杂的问题，我们需要的可能也是一个复杂的处理方案，而复合方案应当就是这样一个值得我们考虑的选择。

参考文献

- 维特根斯坦，2016年《哲学研究》，陈嘉映译，商务印书馆。
- Chalmers, D. J., 1996, *The Conscious Mind*, Oxford University Press.
- Fodor, J., 1974, "Special Science (or The Disunity of Science as a Working Hypothesis) ", in *Synthese* 28 (2) .
- Hyslop, A., 1995, *Other Minds*, Kluwer.
- 2014, "Other Minds", in *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/other-minds>.
- Jacob, P., 2011, "The Direct-Perception Model of Empathy: A Critique", in *Review of Philosophy and Psychology* 2 (3) .
- Krueger, J., 2017, "Direct Social Perception", <https://philpapers.org/archive/KRUDSP.pdf>, Accessed by 2017-07-22.
- Malcolm, N., 1958, "Knowledge of Other Minds", in *Journal of Philosophy* 55 (23) .
- Pargetter, R., 1984, "The Scientific Inference to Other Minds", in *Australasian Journal of Philosophy* 62 (2) .
- Shoemaker, S., 1982, "The Inverted Spectrum", in *The Journal of Philosophy* 79 (7) .
- Spaulding, S., 2015, "On Direct Social Perception", in *Consciousness and Cognition* 36.

(作者单位：厦门大学哲学系)

责任编辑：陈德中

(接上页注①) 对此的回应有二：首先，上述语句“有某种……感受质”中的“有”是什么意思？一个值得考虑的看法是，这里的“有”不是说感受质是某种独立于特定心理语言而存在的形而上学的事例 (entity)，我们只是用特定的心理语言去指称或者来描述它。而是说，感受质乃是认知者Z运用特定心理语言来描述某个 (语言之外的) 事件E的过程中，在Z特定意识活动中的一种“呈现物” (appearance)。因而，这类呈现物不是独立于语言的形而上学上之物，而是特定心理语言表达式的“产物”，不是心理语言之外的，而是心理语言之内的。维特根斯坦亦有类似看法“某种东西是哪一类对象，这由语法来说。” (维特根斯坦，第126页) 其次，即便他者觉得内心中有某种感受质，也不会妨碍我们原则上可以有关于他心的一种认识论上的完整理解。理由就是，如果上文提到的殊型同一原则是合理的，那么不难理解，依据该原则，即便无法把那些心理词汇翻译成认识论词汇，我们也能够从认识论上完整地认识到由特定心理语言所描述的那同一件事 (E)。限于篇幅，不再展开论述。